Integrated Science Education for the 21st Century

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

David Botstein

Lewis-Sigler Institute for Integrative Genomics
Princeton University

Carl Icahn Laboratory-- Rafael Viñoly, architect, 2003

Why is undergradate interest in biology languishing today?

- Depletion of top researchers among active undergraduate teachers.
- Surrogates (TAs, postdocs, even undergraduates) teach introductory courses in math physics chemistry AND Biology; same problem in research labs
- Mature field; perceived need to teach lots of background (history) defers exposure to current day ideas until “later”
- Age-old prerequisites, taught in the same age-old way
- Domination of biology (and general science) teaching by concerns of premedical students
- Widening gap between the effort required of non-science and science students; virtually no students recruited to science

The success rate for interesting quantitatively -oriented students in biology is even more dismal.

The Cultural Gap: Biology and the More Quantitative Sciences

- During the last century, education of biologists has become increasingly less quantitative, to the point that most colleges require minimal mathematics, little or no physics [or physics without calculus (!)], and no computation. Quantitative preparation is usually limited to a number of "prerequisites" that are taken in the freshman and sophomore years, together with pre-medical students. Motivation of both students and faculty in these courses is poor. Today most biology Ph.D. students have only the mathematics and physics they learned in this way.
- This has become an acute problem for biology, because the genome and the computer have transformed biology, and mastery of quantitative tools and theory underlying them is rapidly becoming absolutely essential.
- The cultural gap that has opened has become so wide that effective communication between the biologists and more quantitative scientists has begun actually to inhibit progress. Cross training at the graduate and post-doctoral level, in either direction, has therefore become surprisingly difficult.

These considerations favor an integrated introductory curriculum in which teaching physics, chemistry and biology, together with the relevant mathematics and computational science are taught together

W. Bialek and D. Botstein (2004) Science 303:788

Educating the Biologist of the 21st Century

A curriculum aimed at students considering careers in academic or industrial biology.

- Curriculum aimed at students willing and able to learn **mathematics and computation** at a high level; not intended for all students
- Curriculum intended to **stand beside, and not replace** any existing curricula: standard molecular biology and eco-evo programs that serve premedical students remain in place.
- Curriculum **integrates** suitable math, physics, chemistry and computation.
- Integration of these sciences to begin at the most **introductory level**; no “prerequisites” remain.
- Curriculum features “**project laboratory**” experiences that introduce students to research at the frontier of knowledge early in their studies.
- Advanced studies in math, physics, chemistry and computation are pursued together with majors in these other disciplines.

Getting to a High-Level Integrated Introductory Curriculum

- Accept that integrated curriculum **cannot be the sum** of all the constituent disciplines: choices must be made.

- Study every idea introduced in introductory math, computer science, physics, chemistry and biology: retain those that are **fundamental**, as opposed to merely **traditional.**

- Save the students' time. **Introduce ideas just once**: e.g. calculus & mechanics together; computation and matrix algebra together, etc.

- Provide essential fundamental concepts as required: the "**just in time**" principle. If the scaffold of the course is biology, rigorous introduction of, let us say, Bayes theorem, can and should be done in the context of a genetics problem that requires Bayesian logic. Recognize that "*Learn this now, it is good for you later*" is in reality a form of hazing.

- Teach at a level that **satisfies introductory requirements** in the constituent disciplines wherever possible.

Distinguishing the Fundamental from the Merely Traditional

"First, figure out why you want the students to learn the subject and what you want them to know, and the method will result more or less by common sense".

----Richard Feynman, 1961

There is considerable overhead in teaching the historical/traditional origins of the major ideas. The goal should instead be to teach with the ideas and technology of TODAY.

Examples :

- How do we know DNA is the genetic material?
- How can we measure Avogadro's number?
- How do we find the correct distribution in estimating statistical significance?

In his famous course (and book), Feynman reformulated essentially all of physics, introducing every idea in a non-traditional way.

Advantages of Starting at the Introductory Level

- Interdisciplinary from the beginning: no cultural issues/prejudices (e.g. “math anxiety” or “stamp collecting”) to overcome.
- Can attract undecided (but nevertheless interested) students by the science itself-- no deferral to “prerequisites”.
- Instead, learning of “prerequisite” material (e.g. mathematics and computation) is **motivated by its usefulness and relevance** to the modern and interesting scientific issues (e.g. probability theory motivated by genetics problems)
- Allows for efficient combination of subjects with their related background: enables the “just in time” approach.

It is hard to “go back” to learning prerequisite material if one has advanced in any subject; this is the problem with trying to teach mathematics to biology graduate students or postdocs and vice versa.

Using all the Modern Tools: Especially the Computer

- Simple programming, using modern computer languages (e.g. JAVA, Matlab and/or PERL) are introduced to students from the first. This is easily motivated by problems that involve much iteration or table-parsing to do without any custom programming at all. At Princeton, an introduction to both JAVA & Matlab is included in the new Integrated Introductory Curriculum. Computational problems are prominent on problem sets throughout.

- Algorithms and numerical methods should be introduced and used, often in preference to closed-form solutions. This is the reverse of what is done in most formal courses in physics or statistics, let alone mathematics. Nevertheless, this what actually is done in practice in most modern physics and statistics research.

{repeated observation, in simulations, of the Gaussian turns out to be a great motivator for learning the closed-form mathematics.}

Taking Advantage of Heuristics

Just as algorithms and numerical methods are advantageous in introducing students to quantitative issues in a rigorous way, without first teaching a lot of prerequisite mathematics, **heuristic explanations** should be introduced first, with formal and rigorous calculations deferred. Again, this is **reflective of the real world of research** as opposed to the theoretical world of the formal classroom. Examples:

- Atomic and molecular orbitals
- Ball-and-stick or ribbon models of proteins
- Energy-well diagrams for transition states

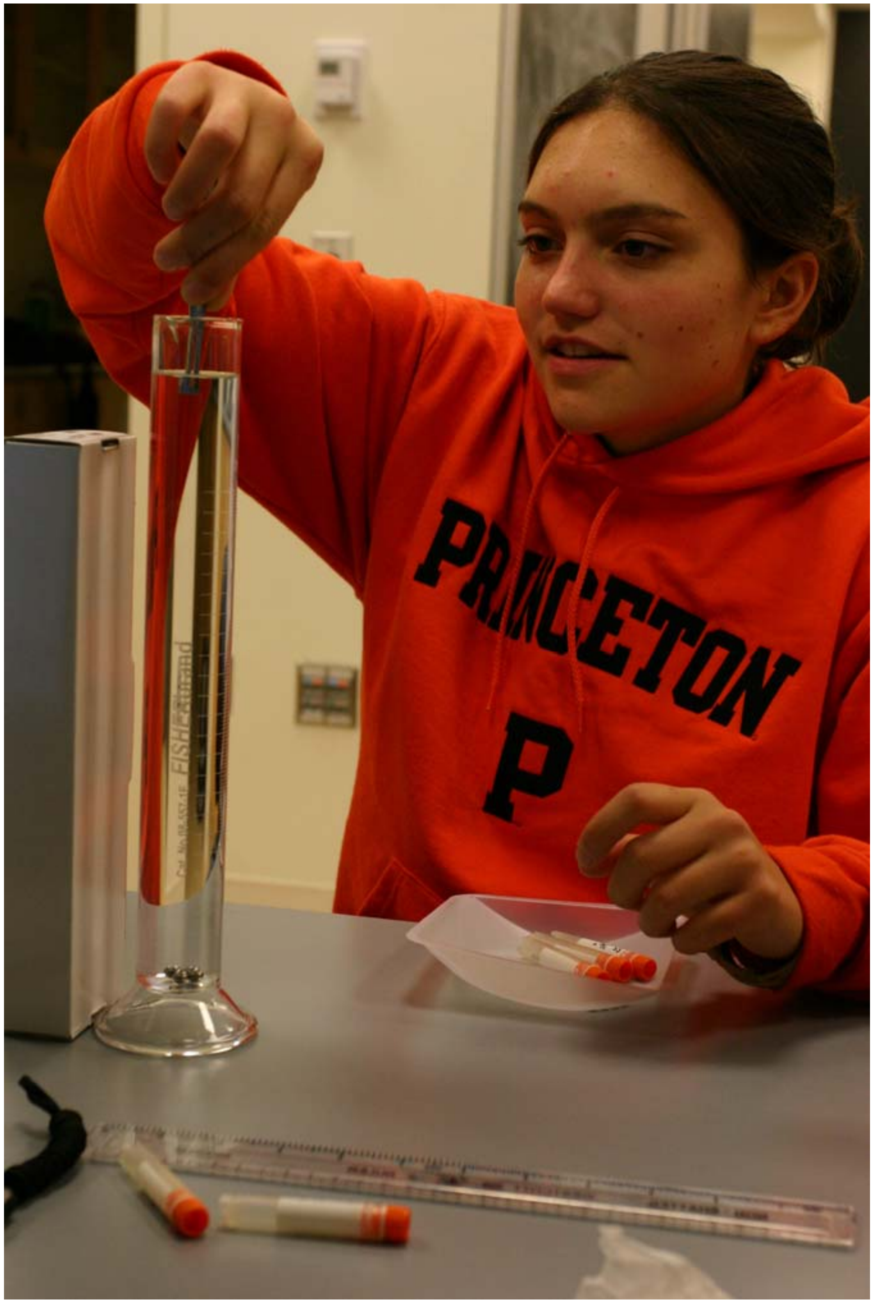
The Princeton Integrated Introductory Science Curriculum
CHM/COS/MOL/PHY 231-236

- Year 1: a double-credit course: 5 h of lecture; 1 3h lab, 1 3h computer lab/section; equivalent to 4 1-semester courses.
- Year 2: two 1-semester courses (no lab); the ensemble will be accepted as preparation for physics, chemistry, CS or biology major.
- Faculty: Bialek & Marlow (Physics), Hecht, Rabinowitz & Groves (Chemistry), Chazelle & Troyanskaya (Computer Science), Botstein & Wieschaus (Molecular Biology); Kruglyak (Ecology & Evolutionary Biology)
- Lewis-Sigler Fellows (Dunham & Ryu) design laboratory sessions
- For students interested in Quantitative Biology or Genomics, while majoring in Chemistry, Physics, Molecular Biology, Ecology & Evolutionary Biology or Computer Science, we will provide Project Laboratory Courses on the MIT model. This takes advantage of Princeton's "certificate program" for inter-disciplinary majors.

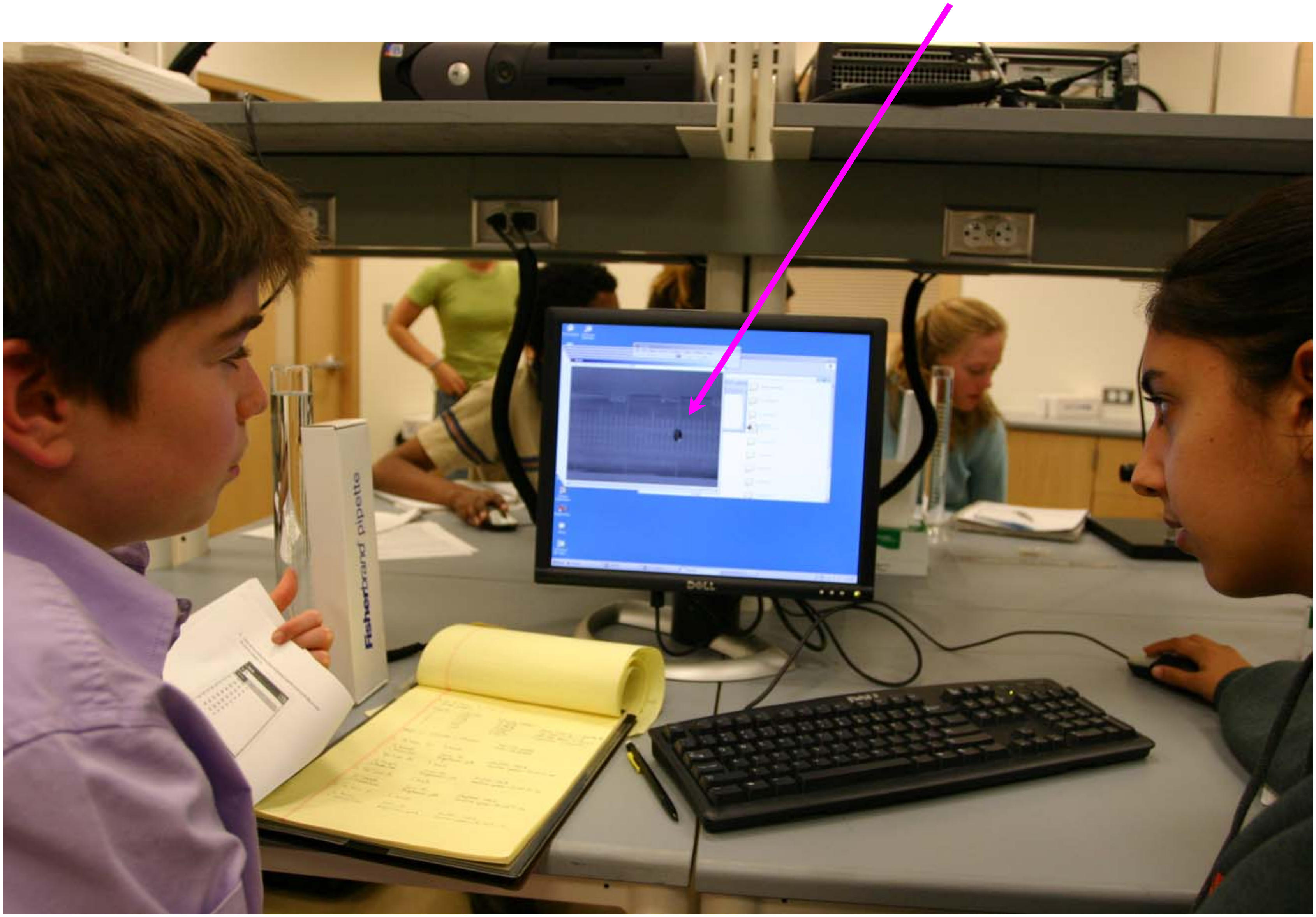
The Princeton Introductory Science Curriculum [a sampling of ideas]

- Integrated approach-- how scientists perceive, measure and model the way things work in the real world. Units include:
  - ---functional relations (linear models)
  - ---dynamical models (differential equations)
  - ---probabilistic models (genetics, molecular motion)
  - ---fields (diffusion, waves, diffraction)
  - ---the quantum world (Schrödinger to the H atom)
  - ---molecular structure (molecular orbitals to reactivity)
- Laboratory exercises-- biology/chemistry experiments, carried out with state-of-the-art technology, producing large data sets:
  - ---Measurement physics: motion through liquids
  - ---Instrumentation: build a photometer using an LED
  - ---Brownian motion: Avogadro & Boltzmann
  - ---Mutation rate by Luria-Delbrück: measurement & simulation
  - ---Synthesis, purification and activity of a simple antibiotic
- Data from the exercises are raw material for the computer lab
- Problem sets: done with analytic or computer methods as appropriate

First Day in the Laboratory: "Theory and Measurement: Moving Through Fluids" (a.k.a. "Life at Low Reynolds Number")


PRINCETON
P


PRINCETON
PRINCETON


fisherbrand pipette
DELL

The first data: a simple measurement, but made with the tools and technology of the 21st Century

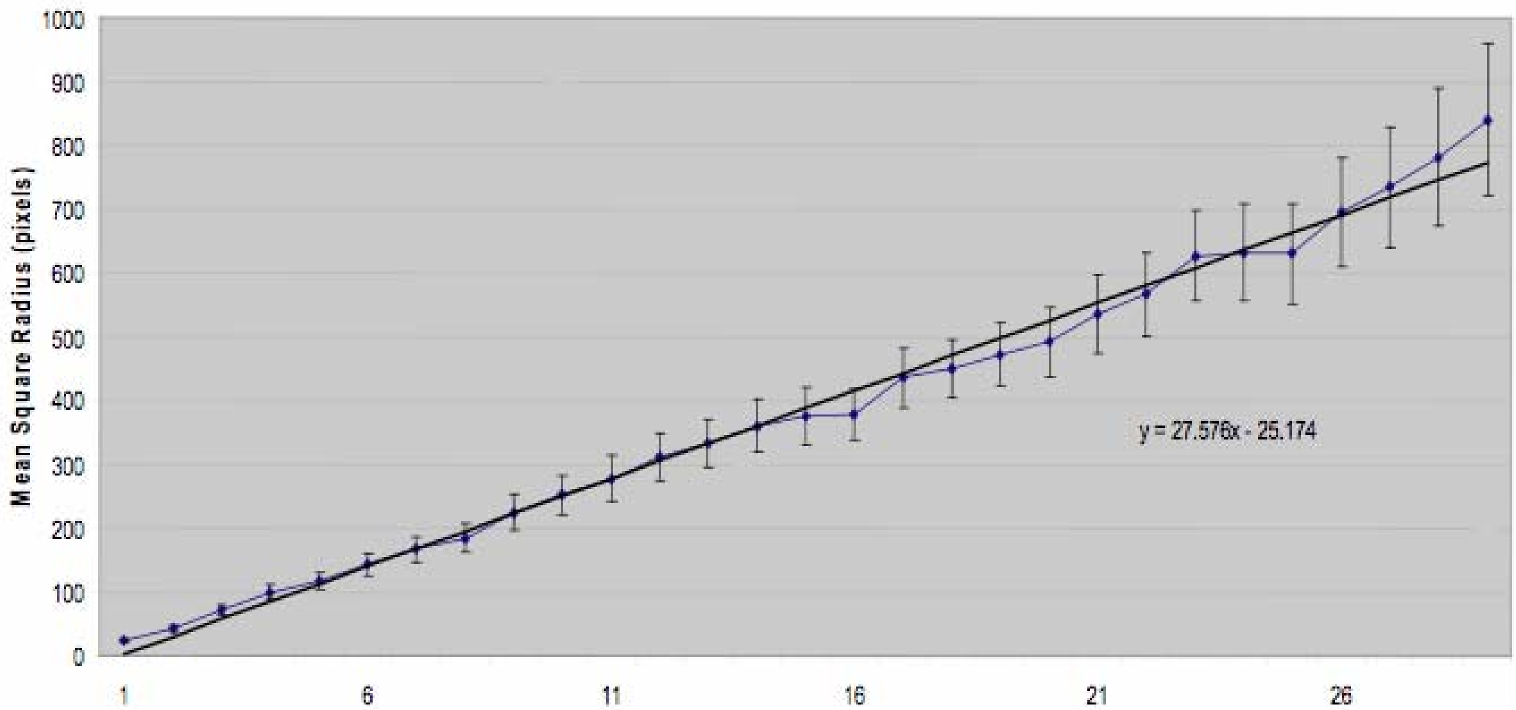
QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

J.P. Owen and W.S. Ryu (2005) *Eur. J. Phys.* **26:**1085–1091

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

Boltzmann's Constant from Brownian Motion

Fluorescent plastic balls imaged in a fluorescence microscope and captured using a firewire camera, ImageJ software and analyzed with MatLab



One of the computer exercises in the spring term of the first year is to produce, in Java, a simulation of Brownian motion in 2 dimensions all the freshmen were able to produce working simulations.

The Second Year: [single course, fall and spring, no lab]

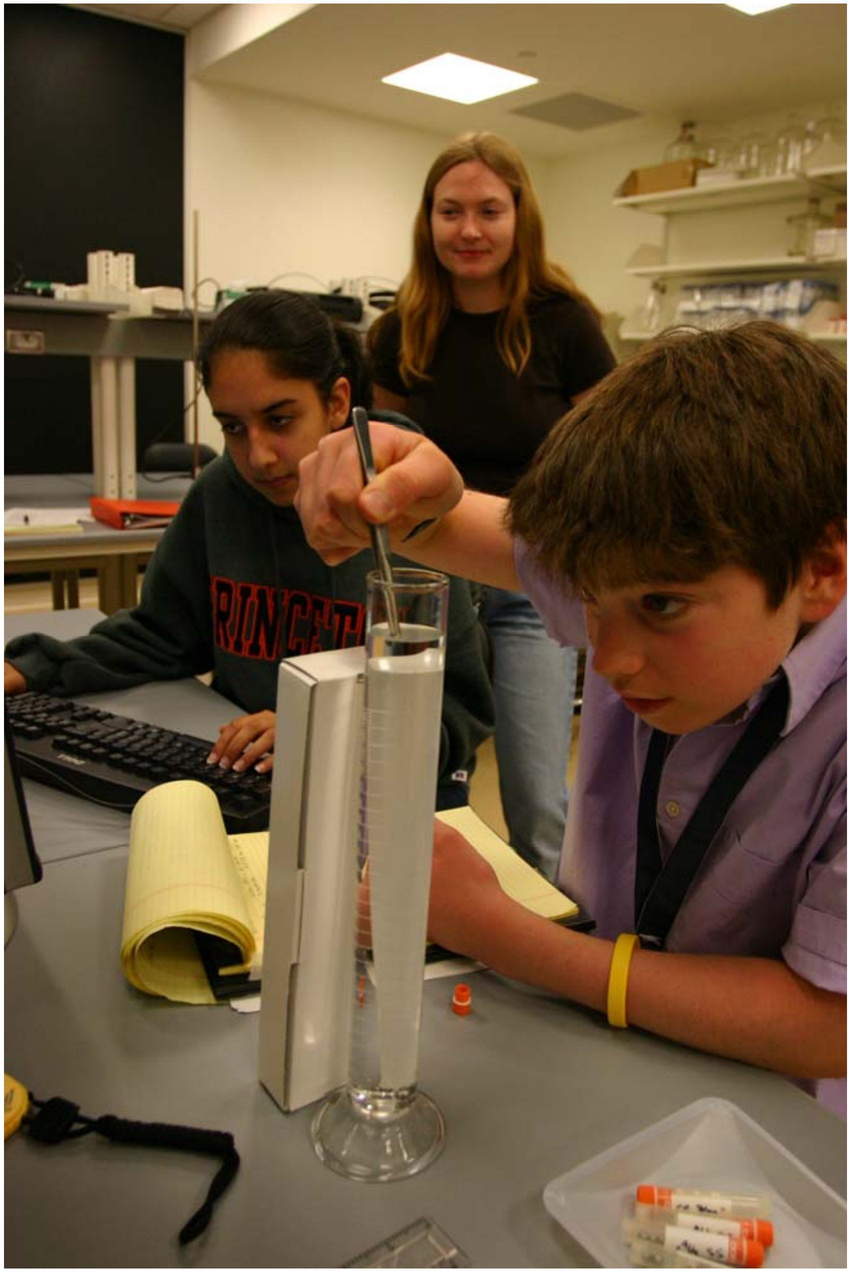
- Shift of emphasis to biology and chemistry, with no change in philosophy: essentials, not traditional, quantitative approach wherever appropriate.
- Covers (and subsitututes for) the standard introductory, genetics and biochemistry courses given by Molecular Biology Department.
  --- Genetics: classical ideas & analysis integrated with genomics, population genetics, genetic epidemiology & molecular evolution
  --- Biochemistry: Integrating basic organic chemistry, metabolism, physiology, structural biology & system-level information.
- Each lecture topic focuses on associated quantitative analysis; problem sets involve computation, including writing of programs.
  --- DNA & RNA: reassociation kinetics [$C_o t$]
  --- Translation: kinetic proofreading
  --- Full treatment of Michaelis-Menten kinetics, Hill coefficients, etc.
  --- Likelihood methods: linkage (LOD scores) & similarity trees

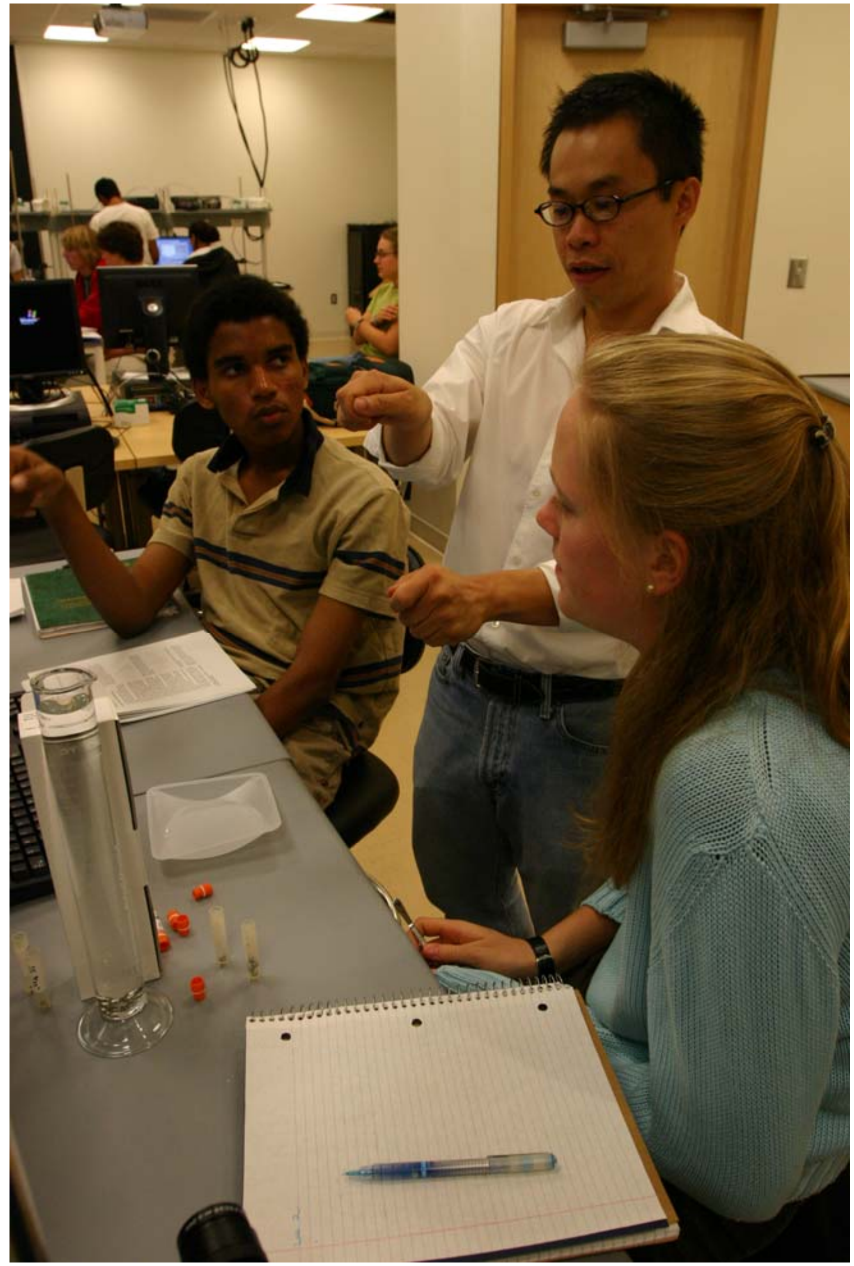
Emerging Principles and Practices

- Focus on the important ideas and take the time to do them well

- Teach every idea and concept from the viewpoint, and with the technology of today:
  - --- introduce underlying concepts “just in time”
  - --- never go out of the way to teach from history
  - --- minimize exposition, detail & “background”,
    if it’s important, do it right, if not, skip it
  - --- coverage is never a goal. Do as much as one can.

- Students enrolled by self-selection
  - --- requires a realistic description of content and level
  - --- provide lots of problem sets and help session
  - --- be flexible: if something doesn’t work, change it
    or do it over again, or provide notes





The Pioneer Class after their Final Exam in Spring, 2005

They appeared for the exam in identical t-shirts of their own design

The T-shirt

…desperately fighting against ignorance,
one integral at a time….

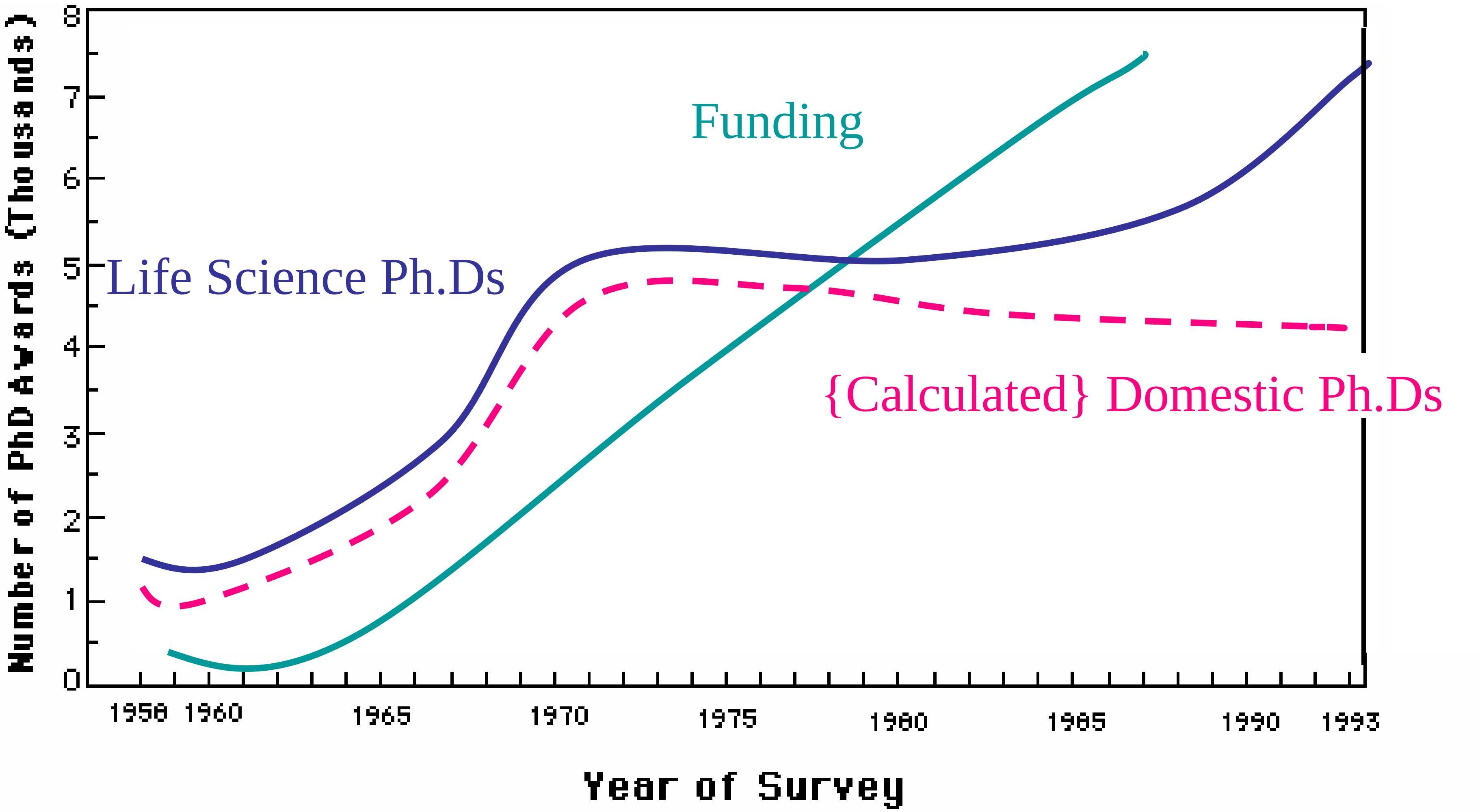
Summary

The turn of the century seems a good time to re-think undergraduate education in biology for students interested in a research career.

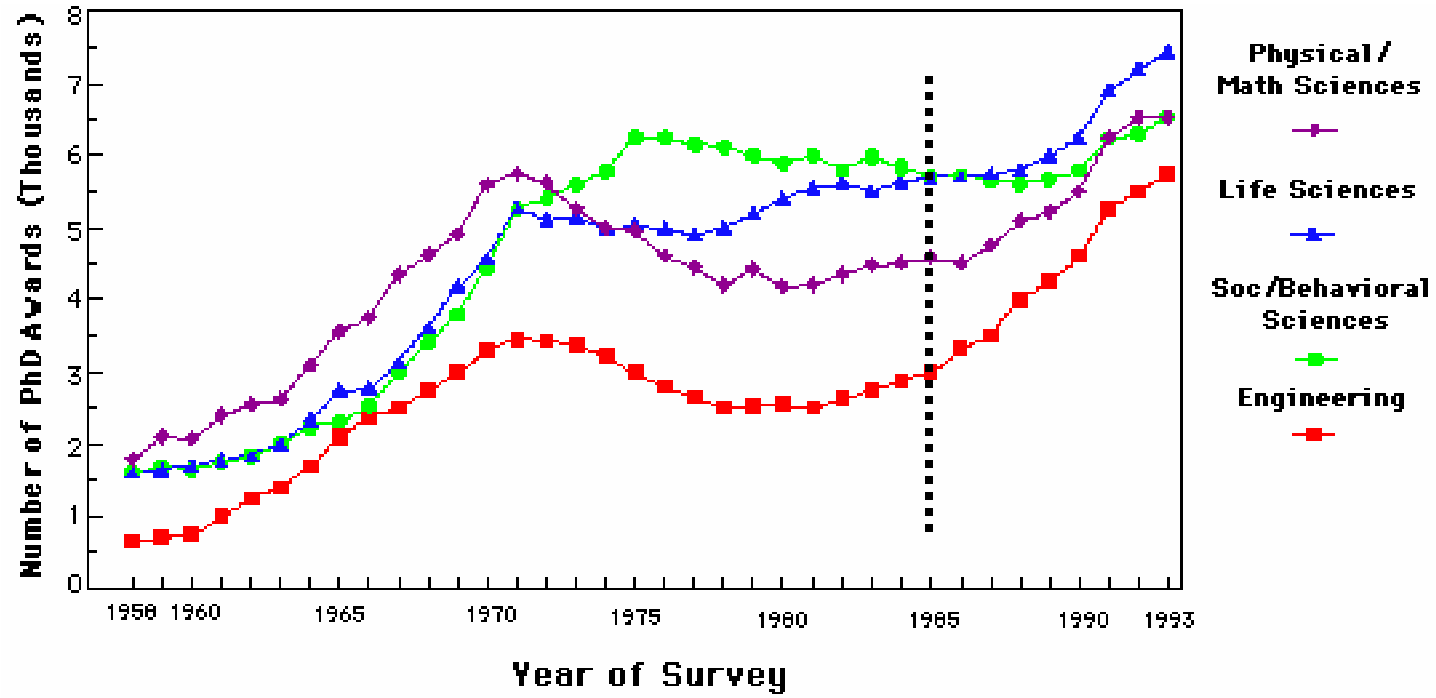
- The **genome** and the **computer** invite a new curriculum that integrates biology with the other sciences.
- A new **curriculum** can produce the new **field** and **culture** needed to attract new students and reversing (at least locally) the dismal trend.
- Teaching should avoid the merely traditional and seek to **illustrate fundamental principles with today's ideas and technology**.
- **Integrated introductory courses**, using computation from the beginning, should prepare students for majors in any science.
- The goal of the curriculum should be to introduce the ideas and techniques of today's science as early as possible, encouraging **authentic original research** in the laboratory and on the computer.
- The **connection between research and teaching** can and should once again be encouraged and rewarded by the universities.

The Diauxie of Basic Science Education 1958-1993

Conclusion: we are failing to interest American undergraduates



Number of doctorates awarded by US institutions 1958-1993



"Most of the net growth after 1985 was due to an increased number of foreign students with temporary student visas"

Source: Reshaping the Graduate Education of Scientists and Engineers COSEPUP/NAS 1996

Two Relevant Reports from the National Academies

[Both available online at www.nap.edu or www.nas.edu]

Reshaping the Graduate Education of Scientists and Engineers

(1995) *National Academy of Sciences, National Academy of Engineering, Institute of Medicine*

Report of COSEPUP study chaired by Phillip A. Griffiths

BIO2010: Transforming Undergraduate Education for Future Research Biologists

(2002) *Report of the Board on Life Sciences* (BLS), chaired by Lubert Stryer

And One from the American Academy of Arts & Sciences

"Science at Liberal Arts Colleges: A Better Education?"

Thomas R. Cech, Daedalus Volume 128, Number 1, Winter 1999

[unfortunately out of print and unavailable]